

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-333898

(P2002-333898A)

(43) 公開日 平成14年11月22日 (2002. 11. 22)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	ターミナル (参考)
G 1 0 L 15/00		A 6 3 F 13/00	F 2 C 0 0 1
A 6 3 F 13/00			M 5 D 0 1 5
		13/12	C
13/12		G 1 0 L 3/00	5 5 1 H
G 1 0 L 15/22			5 6 1 C

審査請求 未請求 請求項の数10 OL (全 9 頁) 最終頁に続く

(21) 出願番号 特願2001-136758 (P2001-136758)

(22) 出願日 平成13年 5 月 7 日 (2001. 5. 7)

(71) 出願人 500299263

株式会社ビバリウム

東京都港区三田 1 丁目 2 番 17 号

(72) 発明者 斉藤 裕

東京都渋谷区神宮前四丁目 11 番 13 号 株式会社ビバリウム内

(72) 発明者 伊藤 成一

東京都渋谷区神宮前四丁目 11 番 13 号 株式会社ビバリウム内

(74) 代理人 100103148

弁理士 山本 輝美

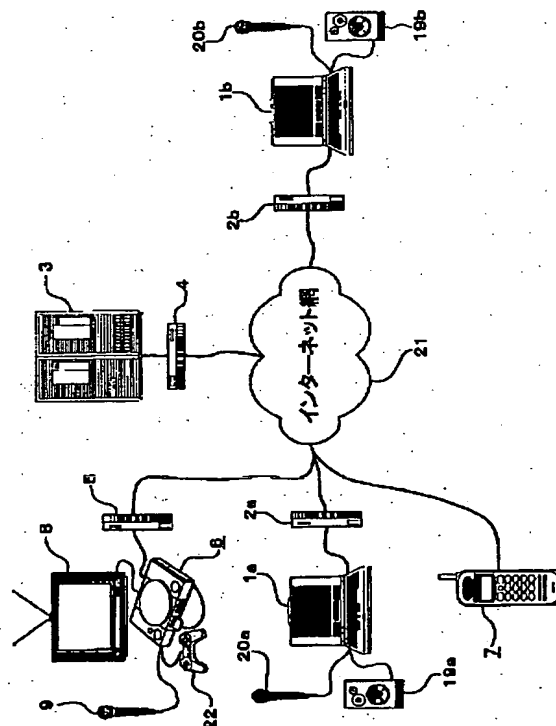
最終頁に続く

(54) 【発明の名称】 電子ペットの音声認識システム

(57) 【要約】

【課題】 本発明は電子ペットの音声認識システムに関し、特に各候補の認識率の差を比較することによって誤認識することなく、且つ効率の良い音声認識を行うことを目的とするものである。

【解決手段】 本発明はシーマン等の電子ペットを使用した電子ペットの音声認識システムであり、電子ペットのプログラムを記憶する端末機器において、人面魚「シーマン」と操作者との会話を行う際、操作者の回答を音声認識し、認識率の異なる複数の候補を提案し、第1候補と第2候補間に所定値以上の認識率の差がある場合第1候補を認識結果とし、第1候補と第2候補間に所定値以上の認識率の差がない場合第1候補に対する確認処理を行い、誤認識を無くすと共に、効率良い音声認識を可能とするものである。



## 【特許請求の範囲】

【請求項1】 複数の端末機器と、ネットワーク回線網を介して接続されデータ通信可能に接続されたサーバコンピュータとを備えた電子ペットの音声認識システムであって、

電子ペットの質問に対して認識率の異なる複数の候補を出力する音声認識手段と、

該音声認識によって得られる第1の候補と第2の候補間の認識率の差を比較する比較手段と、

該比較手段による比較結果が所定値を越えるとき、前記第1の候補を前記音声認識の結果と認定する認定手段と、

前記比較手段による比較結果が所定値を越えないとき、前記第1の候補に対する確認処理を行う確認手段と、を有することを特徴とする電子ペットの音声認識システム。

【請求項2】 前記第1の候補に対する確認処理において、肯定的な応答を受けた場合、該第2候補を前記音声認識の結果と認定することを特徴とする請求項1記載の電子ペットの音声認識システム。

【請求項3】 前記第1の候補に対する確認処理において、否定的な応答を受けた場合、更に前記第2候補に対する確認処理を行うことを特徴とする請求項1記載の電子ペットの音声認識システム。

【請求項4】 前記第2の候補に対する確認処理において、肯定的な応答を受けた場合、該第2の候補を前記音声認識の結果と認定することを特徴とする請求項3記載の電子ペットの音声認識システム。

【請求項5】 前記第2の候補に対する確認処理において、否定的な応答を受けた場合、第3の候補に対する確認処理を行うことを特徴とする請求項3記載の電子ペットの音声認識システム。

【請求項6】 前記第1の候補の認識率は一定値以上であることを特徴とする請求項1、2、3、4、又は5記載の電子ペットの音声認識システム。

【請求項7】 複数の端末機器と、ネットワーク回線網を介して接続されデータ通信可能に接続されたサーバコンピュータとを備えた電子ペットの音声認識方法であって、

電子ペットの質問に対する回答を音声認識する処理と、該音声認識によって複数の認識候補が得られた場合、第1の候補と第2の候補間の認識率の差を比較する比較処理と、

該処理による比較結果が所定値を越えるとき、前記第1の候補を前記音声認識の結果と認定する認定処理と、

前記処理による比較結果が所定値を越えないとき、前記第1の候補に対する確認処理を行う確認処理と、を行うことを特徴とする電子ペットの音声認識方法。

【請求項8】 前記第1の候補の認識率は一定値以上であることを特徴とする請求項7記載の電子ペットの音声

認識方法。

【請求項9】 複数の端末機器と、ネットワーク回線網を介して接続されデータ通信可能に接続されたサーバコンピュータとを備えた電子ペットの音声認識プログラムであって、

電子ペットの質問に対する回答を音声認識する機能と、該音声認識によって複数の認識候補が得られた場合、第1の候補と第2の候補間の認識率の差を比較する比較機能と、

該比較機能による比較結果が所定値を越えるとき、前記第1の候補を前記音声認識の結果と認定する認定機能と、

前記機能による比較結果が所定値を越えないとき、前記第1の候補に対する確認処理を行う確認機能と、を有することを特徴とする電子ペットの音声認識プログラム。

【請求項10】 前記第1の候補の認識率は一定値以上であることを特徴とする請求項9記載の電子ペットの音声認識プログラム。

## 【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、電子ペットの音声認識システム、及びその方法に関する。

【0002】

【従来の技術】近年のエレクトロニクス技術の進歩に伴い、操作者の個性に関する情報を知識として蓄積し、該蓄積した各種データに基づいてその操作者に合致した対応を実現する電子ペットプログラムがある。例えば、水槽の中を泳ぎ回る電子ペットと会話し、操作者とコミュニケーションをとるシステムなどである。

【0003】

【発明が解決しようとする課題】このような電子ペットとのコミュニケーションシステムにおいて、電子ペットとユーザとの会話によってコミュニケーションが進行する。この際、音声認識エンジンや音声合成エンジンを使用して電子ペット及びユーザの音声認識、及び音声合成が行われる。

【0004】しかしながら、従来の音声認識においては、必ずしも確実な認識を行っているとは言えなかった。例えば、人面魚「シーマン」の質問に対する回答を認識する場合、最も認識率の高い認識結果を回答としていた。このため、それほど認識率自体が高くない場合でも回答であると判断し、誤認識を行う危険があった。

【0005】また、必ず認識結果をユーザに確認する処理を含めるシステムもある。しかし、この場合には極めて効率の悪い音声認識方法となる。本発明は上記課題に鑑み、認識率の差を比較することによってより正確な音声認識を行うことを目的とするものである。

【0006】

【課題を解決するための手段】上記課題は請求項1記載

の発明によれば、複数の端末機器と、ネットワーク回線網を介して接続されデータ通信可能に接続されたサーバコンピュータとを備えた電子ペットの音声認識システムであって、電子ペットの質問に対して認識率の異なる複数の候補を出力する音声認識手段と、該音声認識によって得られる第1の候補と第2の候補間の認識率の差を比較する比較手段と、該比較手段による比較結果が所定値を越えるとき、前記第1の候補を前記音声認識の結果と認定する認定手段と、前記比較手段による比較結果が所定値を越えないとき、前記第1の候補に対する確認処理を行う確認手段とを有する電子ペットの音声認識システムを提供することによって達成できる。

【0007】このように構成することによって、第1候補と第2候補に大きな認識率の差があるとき、第1候補を正しい回答であると判断して音声認識し、効率良い音声認識処理を行うものである。

【0008】請求項2の記載は、上記請求項1記載の発明において、前記第1の候補に対する確認処理において、肯定的な回答を受けた場合、該第1候補を前記音声認識の結果と認定する構成である。

【0009】このように構成することによって、前記第1候補を音声認識の結果と判断することができ、誤認識を防止することができる。請求項3の記載は、上記請求項1記載の発明において、前記第1の候補に対する確認処理において、否定的な応答を受けた場合、更に前記第2候補に対する確認処理を行う構成である。

【0010】このように構成することによって、第2候補以下についても効率良く、且つ誤認識を行うことなく、音声認識処理を行うことが可能となる。請求項4の記載は、上記請求項3の記載において、前記第2の候補に対する確認処理において、肯定的な応答を受けた場合、該第2の候補を前記音声認識の結果と認定する構成である。

【0011】このように構成することによっても、第2候補以下についても効率良く、且つ誤認識を行うことなく、音声認識処理を行うことが可能となる。請求項5の記載は、上記請求項3の記載において、前記第2の候補に対する確認処理において、否定的な応答を受けた場合、第3の候補に対する確認処理を行う構成である。

【0012】請求項6の記載は、上記請求項1乃至5の記載において、前記第1の候補の認識率は一定値以上である。このように構成することにより、例えば第1の候補が低い認識率であり、しかも第2の候補との間に所定値以上の認識率の差がある場合、低い認識率の第1候補が回答であると判断されることを防止する。

【0013】上記課題は請求項7記載の発明によれば、複数の端末機器と、ネットワーク回線網を介して接続されデータ通信可能に接続されたサーバコンピュータとを備えた電子ペットの音声認識方法であって、電子ペットの質問に対する回答を音声認識する処理と、該音声認識

によって複数の認識候補が得られた場合、第1の候補と第2の候補間の認識率の差を比較する比較処理と、該処理による比較結果が所定値を越えるとき、前記第1の候補を前記音声認識の結果と認定する認定処理と、前記処理による比較結果が所定値を越えないとき、前記第1の候補に対する確認処理を行う確認処理とを行う電子ペットの音声認識方法を提供することによって達成できる。

【0014】本発明は方法の発明であり、このように構成することによっても、第1候補と第2候補に大きな認識率の差があるとき、第1候補を正しい回答であると判断し、音声認識することができる。

【0015】請求項8の記載は、上記請求項7記載の発明において、前記第1の候補の認識率は一定値以上である。上記課題は請求項9記載の発明によれば、複数の端末機器と、ネットワーク回線網を介して接続されデータ通信可能に接続されたサーバコンピュータとを備えた電子ペットの音声認識プログラムであって、電子ペットの質問に対する回答を音声認識する機能と、該音声認識によって複数の認識候補が得られた場合、第1の候補と第2の候補間の認識率の差を比較する比較機能と、該比較機能による比較結果が所定値を越えるとき、前記第1の候補を前記音声認識の結果と認定する認定機能と、前記機能による比較結果が所定値を越えないとき、前記第1の候補に対する確認処理を行う確認機能とを有する電子ペットの音声認識プログラムを提供することによって達成できる。

【0016】本発明はプログラムの発明であり、このように構成することによっても、第1候補と第2候補に大きな認識率の差があるとき、第1候補を正しい回答であると判断し、音声認識することができる。

【0017】請求項10の記載は、上記請求項9記載の発明において、前記第1の候補の認識率は一定値以上である。

【0018】

【発明の実施の形態】以下、本発明の実施形態を図面に基づいて説明する。図1は、本発明の電子ペットを利用する音声認識システムのシステム構成図である。本例のシステムは、同図に示すように、利用者が所有する複数のプラットフォームであるパーソナルコンピュータ1a、及び1bと、コンピュータゲーム機6と、携帯電話7と、電子ペットである人面魚「シーマン」の提供サービスを行うサービス事業者が所有するサーバコンピュータ3とから構成されている。

【0019】上記パーソナルコンピュータ1a、1b、並びにゲーム機6は各通信装置2a、2b、及び5を介してそれぞれインターネット21に接続され、また携帯電話7は図示しない中継局や中央制御局を介して、インターネット21に接続されている。

【0020】また、インターネット21に接続されたパーソナルコンピュータは1a、1bで示すが、上記2台

のコンピュータ以外に、多数のコンピュータがインターネット21に接続されている。尚、パーソナルコンピュータ1aは、例えばOSとしてウインドウズを使用し、パーソナルコンピュータ1bは、OSとして例えばマッキントッシュを使用する。

【0021】 先ず、本例において受信者並びに発信者が使用するパーソナルコンピュータ1の構成を図2に示す。尚、パーソナルコンピュータ1の構成説明において、代表してパーソナルコンピュータ1aの例で説明する。

【0022】 図2に示すように、パーソナルコンピュータ1aの内部には、データの送受を行うデータバス10が配設され、このデータバス10に中央処理装置（以下、CPUで示す）11や、RAM12、リアルタイムクロック（RTC）17等が接続されている。

【0023】 CPU11は後述する利用者への質問の出題や回答の受付を音声合成により表示装置に表示される電子ペットの言葉として出力する出題処理や、利用者の回答を音声認識して予め登録されている複数の回答候補から該当する項目を選出登録して利用者の個性情報として蓄積する回答受付処理等を実施する。この処理の際、ワークエリアとしてRAM12を使用する。

【0024】 また、リアルタイムクロック（RTC）17は、例えばデータ更新の日時情報に使用される現在の時刻情報や、任意の年月日の曜日等のカレンダー情報を出力する。また、入力装置16はキーボードやマウス等であり、キーボードやマウス等の操作情報をCPU11に通知する。

【0025】 表示装置14はCRT又はLCD等のディスプレイであり、後述するシーマンの表示を行う。また、通信インターフェイス13は前述の通信装置2（2a）に接続され、通信装置2（2a）及びインターネット21を介してサーバコンピュータ3との間でデータの送受信を行う。

【0026】 また、音声入出力装置18には外部機器であるスピーカ19やマイク20に接続され、A/D・D/Aコンバータを有する。音声入出力装置18は、後述する音声合成プログラムにより生成された音声データをアナログの音声と変換（D/A変換）し、上記スピーカ19に出力すると共に、上記マイク20から入力された音声をデジタルデータに変換（A/D変換）して出力する。

【0027】 記憶装置15は、磁気ディスクや光磁気ディスクから成り、上記CPU11の制御に従ってデータやプログラムの書き込み、読み出し処理が行われる。この記憶装置15には前述の表示装置14に表示される電子ペットである人面魚「シーマン」の画像や動作処理等が記述された電子ペットプログラムや、前記人面魚「シーマン」の声をテキストデータに基づいて音声出力するための音声合成プログラムや、発信者或いは受信者の入

力音声をテキストデータに変換するための音声認識プログラム等が記憶されている。

【0028】 また、記憶装置15には人面魚「シーマン」がアプリケーションにおいて使用し、理解する単語群が「辞書」として登録されている。図3はこの辞書の構成を示す図であり、カテゴリ毎に登録されている。

【0029】 例えば、項目番号1には人面魚「シーマン」の質問「おまえは男」、及び当該質問に対する回答a、bが登録されている。また、項目番号2には人面魚「シーマン」の質問「おまえの歳は」、及び当該質問に対する回答a、b・・・が登録されている。以下、同図に示す通りであり、人面魚「シーマン」の質問、及びその回答が複数単語辞書として登録されている。そして、後述する認識処理の際、ユーザの回答に対して各単語毎に比較処理を行い、各単語毎に認識率を出力する。

【0030】 上記構成の音声認識システムにおいて、以下に処理動作を説明する。図4は本例の処理動作を説明するフローチャートである。先ず、人面魚「シーマン」からの質問が行われる（ステップ（以下、STで示す）1がYES）。この質問は前述のデータベースに記憶されたデータ順に行われ、この質問は上記パーソナルコンピュータ1aのスピーカ19aから流れる（ST2）。例えば、人面魚「シーマン」が行う質問が、前述の図3に示す項目番号2の場合、「おまえの歳は」の質問である。この質問はスピーカ19aから流れ、ユーザは質問を理解する。

【0031】 次に、ユーザは上記質問に答えて、マイク20aに向かって回答を行う（ST3がYES）。CPU11は上記回答から音声認識を行い、複数の候補を選択する（ST4）。例えば、ユーザの回答が「27歳」である場合、図5に示す候補が選択される。すなわち、この場合の回答は上記項目番号2に対応する回答a、b、c、・・・の中から選択され、各単語に対する音声比較、例えば積分値の比較等から認識率の高い順に出力される。例えば、図5に示すように、第1候補として「じゅうなさい」、第2候補として「にじゅうなさい」、第3候補として「ごじゅうなさい」、・・・が出力される。

【0032】 また、同時に各候補の認識率も表示される。例えば、第1候補が60%であり、第2候補が55%であり、第3候補が45%である。次に、CPU11は上記第1候補と第2候補の認識率の差がnポイント以下か判断する（ST5）。このnポイントは予め設定されており、例えば経験上誤認識を起こさない値である。ここで、例えば上記nポイントが「20」に設定されていれば、上記図5の例ではYESである。

【0033】 このように、第1候補と第2候補の認識率の差がnポイント以下であれば（ST5がYES）、第1候補をユーザに提示する（ST6）。すなわち、この場合認識結果が確実ではないので、ユーザに確認を促

す。例えば、上記例では「今、17歳って言った」という確認を促す(図6参照)。一方、前述の判断(ST5)において、第1候補と第2候補との差がnポイント以上であれば(ST5がNO)、第1候補に確定し、回答取得を行うと共に、上記確認処理を行うことなく、例えば人面魚「シーマン」はスピーカ19aから「17歳か」と言う(図7参照)。

【0034】次に、上記第1候補の確認処理(ST6)を行い、スピーカ19aからの報音と図6に示す表示を行った結果、ユーザから肯定的な回答があれば、この場合にも第1候補を回答として確定する(ST8がYES、ST9)。例えば、ユーザがマイク20aに向かって「はい」、「うん」、「そう」等の答えを返してきた場合、第1候補を回答として確定する。

【0035】一方、ユーザが「ちがう」、「いいえ」等の否定的な回答を行った場合、人面魚「シーマン」は次の候補があるか判断し、第2候補がある場合、第2候補を提示する(ST10がYES、ST11)。例えば、図5に示す例の場合、第2候補である「にじゅうななさい」の提示を行う。すなわち、人面魚「シーマン」は「じゃ、27歳って言った」という質問をする。

【0036】以下、同様に処理を行い、第2候補以降の提示に対して肯定的な回答があれば、当該回答を回答取得とする。例えば、第2候補の27歳の提示に対してユーザから肯定的な答えがあれば、ユーザの年齢は27歳であると分かり、第3候補の57歳の提示に対して肯定的な答えがあれば、ユーザの年齢は57歳であると分かる。

【0037】以上の処理を繰り返すことによって、人面魚「シーマン」はユーザの正確な年齢を知ることができる。尚、全ての候補がユーザによって否定された場合、最初の質問に会話を戻す。

【0038】以上のように制御することによって、人面魚「シーマン」の質問に対して音声認識の認識率を出力し、該認識率に基づいて音声認識処理を進め、ストレスのないユーザとの対話を実現するものである。

【0039】尚、人面魚「シーマン」の質問はユーザに対する年齢に対する質問であったが、年齢以外の質問であっても同様の処理によって、ユーザからの正確な情報を得ることができる。

【0040】また、第1候補の認識率が一定値以下の場合、例えば第1候補と第2候補間に所定値以上の認識率の差があったとしても、音声認識結果とすることなく、かかる場合には両候補について確認処理を行う等の手続きを行い、誤認識のないシステムとする。

【0041】尚、上記処理において、使用する端末機器はパーソナルコンピュータ1aであったが、他のパーソ

ナルコンピュータ1bであってもよく、又はゲーム機6や携帯電話7であっても良い。

【0042】また、パーソナルコンピュータ等はサーバに接続されていない状態であってもよく、対応する機能を回路やプログラムによって単独に保有する構成であればよい。

#### 【0043】

【発明の効果】以上説明したように、本発明の電子ペットの音声認識システムによれば、誤認識を無くすと共に、効率良い音声認識を可能とするものである。

【0044】また、第1候補の認識率が一定値以下の場合、両候補について確認処理を行う等の処理により、より正確な音声認識システムとすることができる。

#### 【図面の簡単な説明】

【図1】本発明の電子ペットの音声認識システムのシステム図である。

【図2】本発明の実施例において用いた利用者が所有するコンピュータを示すブロック図である。

【図3】記憶装置に登録されたデータベースの構成を示す図である。

【図4】本例における処理動作を説明するフローチャートである。

【図5】第1候補、第2候補、第3候補の認識率を示す図である。

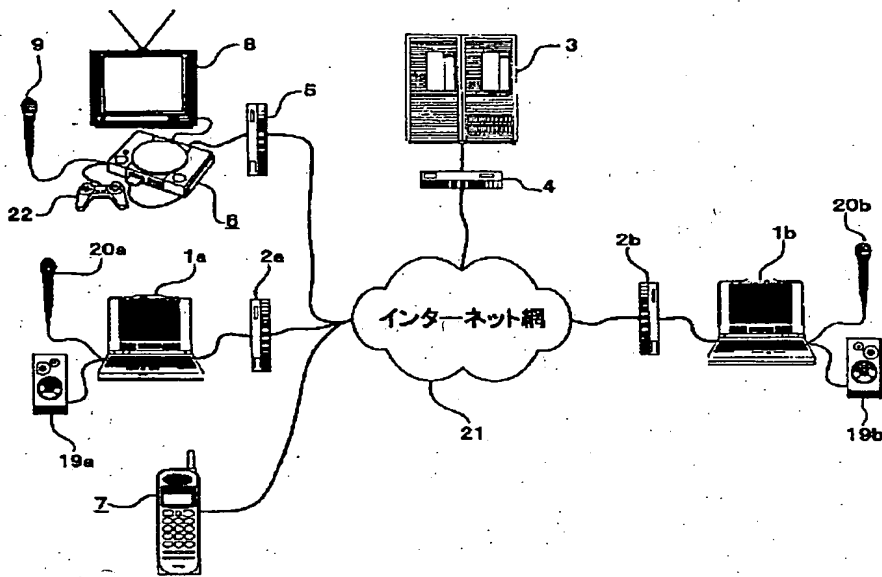
【図6】人面魚「シーマン」の発音状態を表示する図である。

【図7】人面魚「シーマン」の発音状態を表示する図である。

#### 【符号の説明】

- |    |                  |
|----|------------------|
| 1  | パーソナルコンピュータ(利用者) |
| 2  | 通信装置             |
| 3  | サーバコンピュータ        |
| 4  | 通信装置             |
| 5  | 通信装置             |
| 6  | コンピュータゲーム機       |
| 7  | 携帯電話             |
| 8  | テレビ              |
| 9  | マイク              |
| 10 | データベース           |
| 11 | 中央精算処理装置(CPU)    |
| 12 | RAM              |
| 13 | 通信インターフェイス       |
| 14 | 表示装置             |
| 15 | 記憶装置             |
| 19 | スピーカ             |
| 20 | マイク              |
| 21 | インターネット網         |

【図1】



【図5】

例)

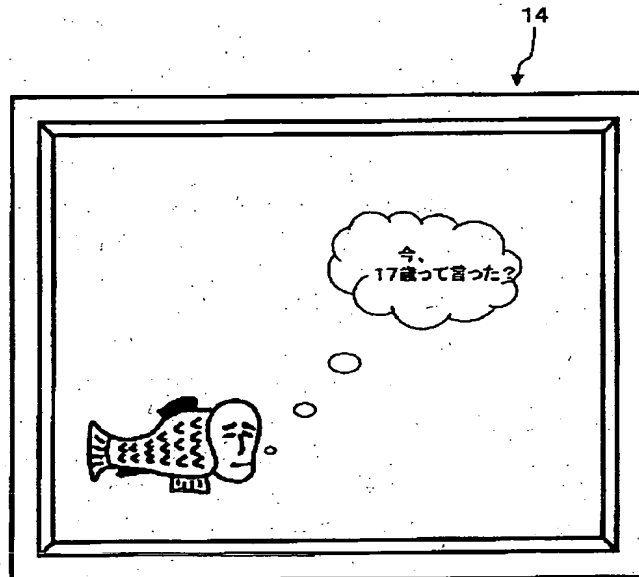
認識率

- 第1候補 : じゅうななさい : 60  
 第2候補 : にじゅうななさい : 55  
 第3候補 : ごじゅうななさい : 45

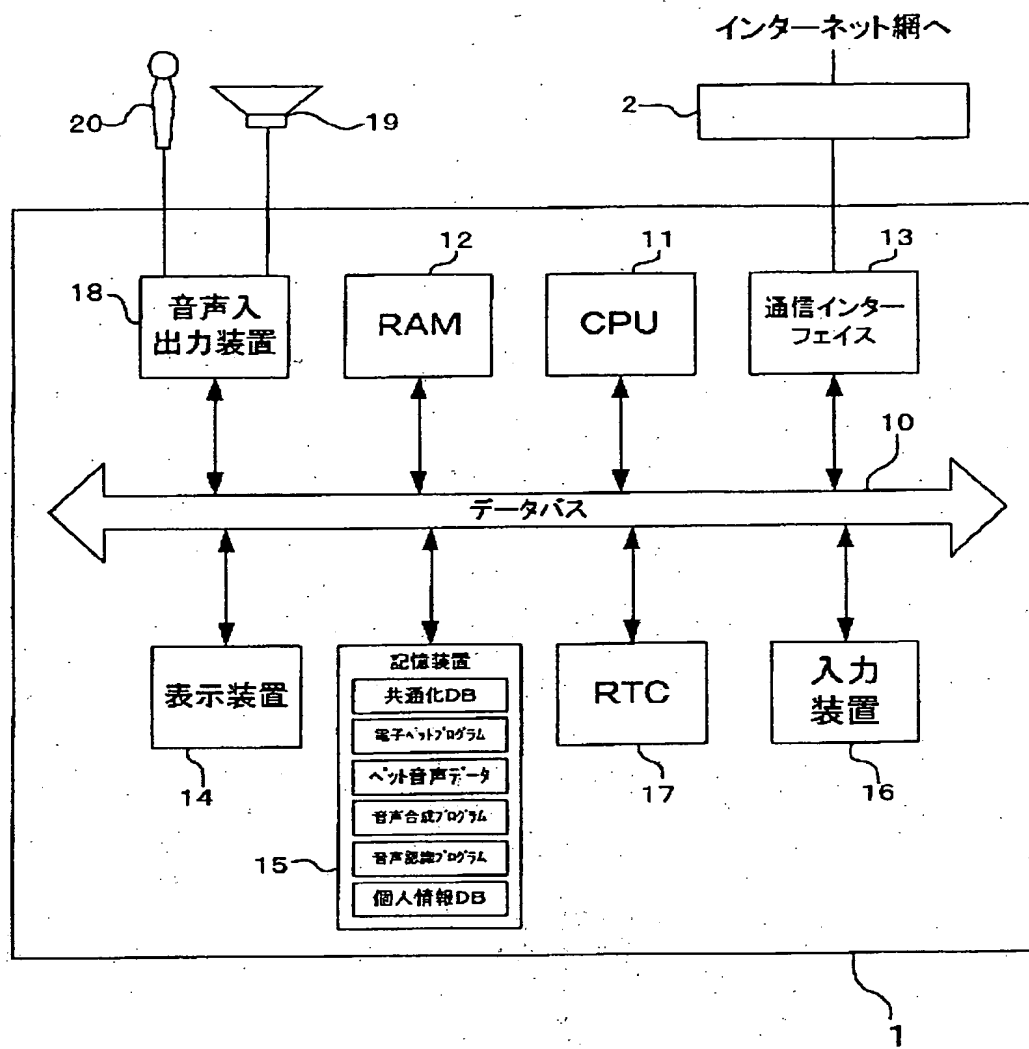
【図3】

質問回答データベース		
項目番号	質問内容	回答候補
1	おまえは男か?	a はい(そうだ、ああ)
		b いいえ(違う、いやだ~)
2	おまえの歳は?	a 10歳
		b 11歳
		c 12歳
		d 13歳
		e 14歳
N	今、欲しいものはあるの?	a ある
		b ない
		c 判らない
		d ものではない
N'	それは何?	単語登録
N+1	今、彼女(氏)はいるの?	a いる
		b いない
		c 判らない
N+2	付き合いはどの位?	a 3ヶ月未満
		b 1年来満

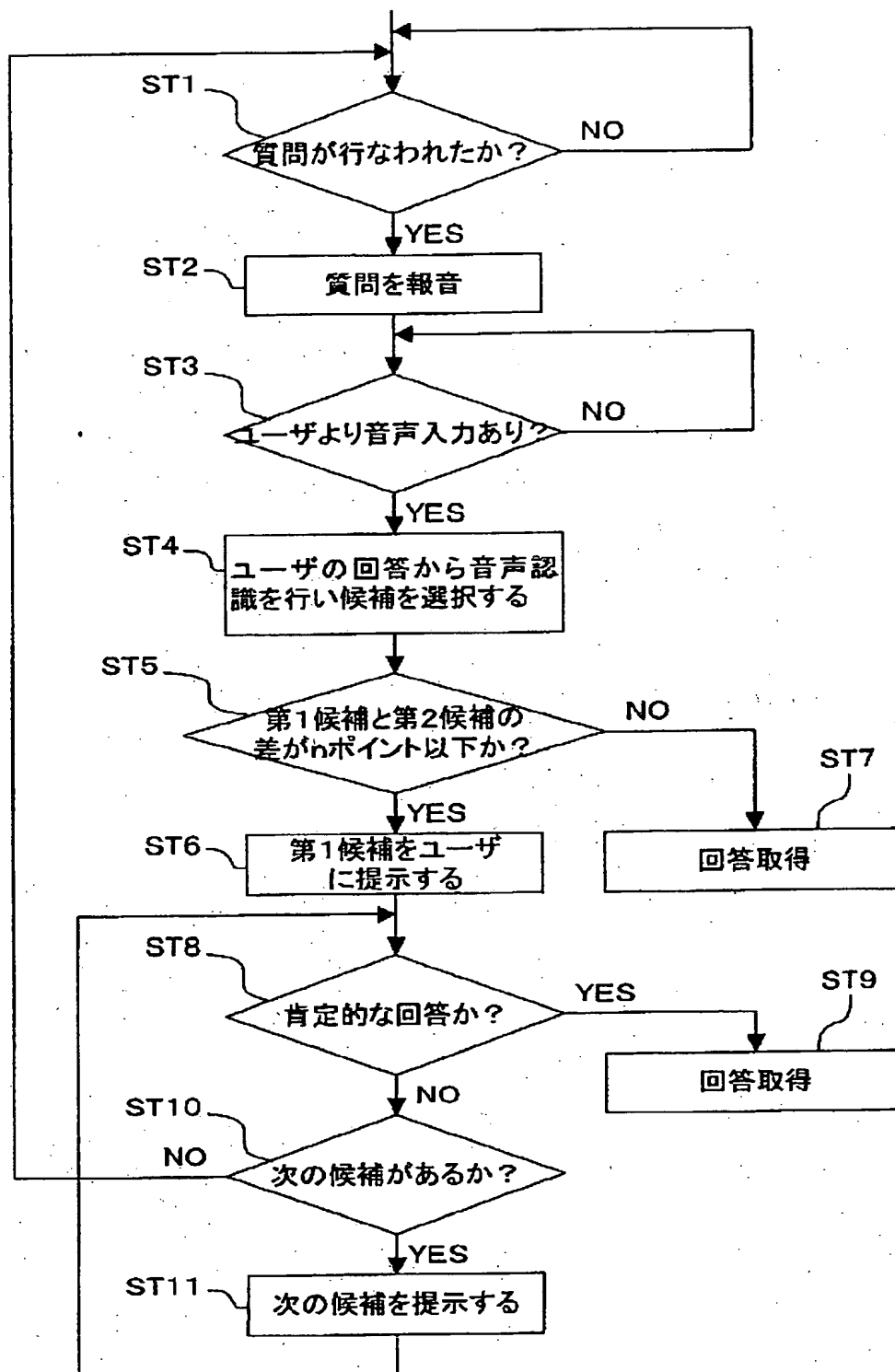
【図6】



【図2】

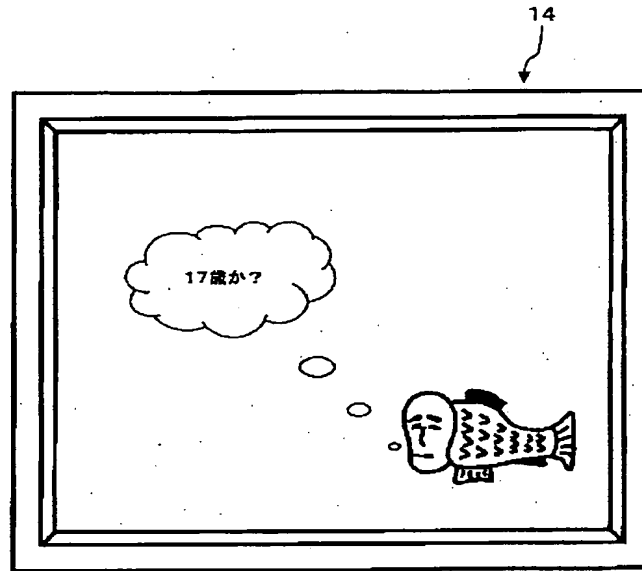


【図4】





【図7】



フロントページの続き

(51) Int. Cl.<sup>7</sup>

識別記号

F I  
G 1 0 L 3/00

フィールド (参考)

5 7 1 U  
5 6 1 E

F ターム (参考) 2C001 AA00 AA11 BA00 BA06 BA07  
CA00 CA07 CB01 CB04 CB08  
CC01 CC08  
5D015 KK02 LL04 LL05

This Page is inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images problems checked, please do not report the problems to the IFW Image Problem Mailbox**